



SesIndexCreatoR : Un package R pour la création et la visualisation d'indices socioéconomiques

Benoît Lalloué, Séverine Deguen, Jean-Marie Monnez, Cindy M. Padilla,
Wahida Kihal, Denis Zmirou-Navier, Nolwenn Le Meur

► To cite this version:

Benoît Lalloué, Séverine Deguen, Jean-Marie Monnez, Cindy M. Padilla, Wahida Kihal, et al.. SesIndexCreatoR : Un package R pour la création et la visualisation d'indices socioéconomiques. Deuxièmes rencontres R, Jun 2013, Lyon, France. hal-01433087

HAL Id: hal-01433087

<https://hal.science/hal-01433087>

Submitted on 12 Jan 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SesIndexCreator: Un package R pour la création et la visualisation d'indices socioéconomiques

**B. Lalloué^{a,b,c}, S. Deguen^{a,b}, J.-M. Monnez^c, C. Padilla^{a,b},
W. Kihal^a, D. Zmirou-Navier^{a,b,d} and N. Le Meur^{a,e}**

^aEHESP Rennes, Sorbonne Paris Cité
Rennes, France

^bInserm, UMR IRSET Institut de recherche sur la santé l'environnement et le travail - 1085
Rennes, France

^cIECL, Institut Elie Cartan Nancy, CNRS : UMR 7502
Université de Lorraine, INRIA BIGS, France

^dFaculté de Médecine

Université de Lorraine, France

^eUMR936 INSERM, Université de Rennes 1
Rennes, France

benoit.lalloue@ehesp.fr ; severine.deguen@ehesp.fr ; jean-marie.monnez@univ-lorraine.fr ;
cindy.padilla@ehesp.fr ; wahida.kihal@ehesp.fr ; denis.zmirou@ehesp.fr ;
nolwenn.lemeur-rouillard@ehesp.fr

Mots clefs : Analyse en composantes principales, Classification Ascendante Hiérarchique, Statut socio-économique.

Afin de prendre en compte les inégalités sociales de santé, il est fréquent d'utiliser des indices socio-économique qui synthétisent différents aspects du statut socio-économique (SES) à l'échelle de l'individu, du voisinage ou de la région. De nombreux indices existent déjà mais nombre d'entre eux utilisent un faible nombre de variables, des méthodes d'agrégation simples et/ou sélectionnent les variables uniquement depuis la littérature. Nous avons développé une nouvelle procédure de création d'indices SES [1] visant à sélectionner à partir d'un grand nombre de variables celles qui composeront l'indice et à les agréger à l'aide de techniques d'analyses de données. Cette procédure est plus complexe à mettre en oeuvre pour des non statisticiens que d'autres approches existantes, c'est pourquoi nous l'avons implémentée dans un package R nommé SesIndexCreator. Ce package vise à donner des outils aussi simples d'emploi que possible pour effectuer la procédure de création et exploiter ses résultats, tout en conservant les différentes possibilités et la flexibilité qu'elle offre.

Le package SesIndexCreator dépend des packages FactoMineR et class. En particulier, la plupart des fonctions d'analyse de données et de visualisation sont issues et adaptées du package FactoMineR. Les sources du package SesIndexCreator sont disponibles librement sur http://www.equitarea.org/documents/packages_1.0-0/. Le package est composé de trois fonctions principales, de leurs fonctions de visualisation et d'impression associées, et de plusieurs autres fonctions internes :

- La fonction SesIndex crée un indice socio-économique utilisant la procédure décrite ailleurs [1]. Il est possible de choisir l'ensemble de variables de départ, les potentielles variables "redondantes" (représentant une même notion socio-économique), les potentielles unités illustratives, la méthode de sélection (ACP ou AFM) ou l'étape de la procédure à réaliser.

Les résultats incluent l'indice final et tous les résultats des étapes intermédiaires.

- La fonction `SesClassif` crée des catégories socio-économiques basées sur un indice créé par la fonction `SesIndex`, en offrant la possibilité d'utiliser différentes techniques : CAH (avec consolidation par les k plus proches voisins ou non), quantiles ou intervalles de même taille. Les résultats incluent à la fois le jeu de données d'origine auquel a été ajouté la catégorie de chaque individu, ainsi que les résultats de la technique de classification.
- La fonction `SesReport` crée un fichier HTML contenant un rapport synthétisant les résultats des différentes étapes de création d'un indice effectuée avec la fonction `SesIndex` et, s'il y en a une, de la classification de l'indice faite avec la fonction `SesClassif`. La fonction `SesReport` permet aussi l'export d'un fichier CSV contenant les données d'origine, l'indice calculé et l'éventuelle classification.

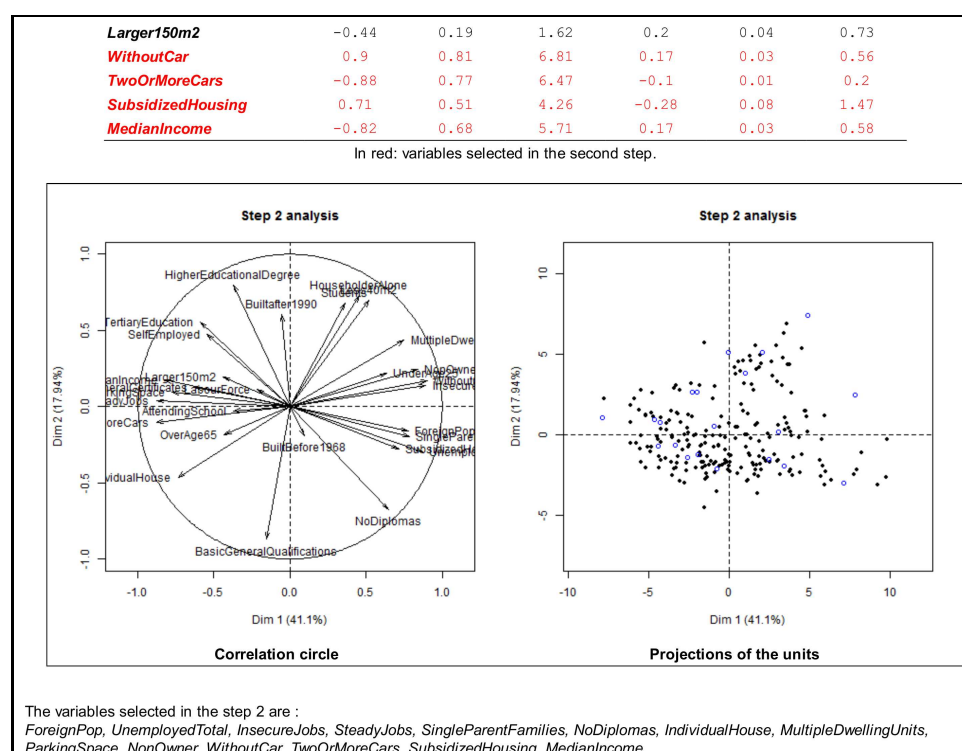


Figure 1: Extrait d'un rapport d'exemple généré par la fonction `SesReport`

Le package `SesIndexCreator` propose ainsi sous une forme "tout en un" l'ensemble des outils nécessaires pour appliquer, interpréter et utiliser la procédure de création d'un indice socio-économique développée, y compris pour des utilisateurs non statisticiens. Nous projetons par ailleurs d'étendre le package dans le futur en ajoutant entre autres de nouvelles méthodes de classification, davantage d'outils d'interprétation ou encore d'autres moyens de visualisation (comme de la cartographie).

Références

- [1] Lalloué B, Monnez JM, Padilla C, Kihal W, Le Meur N, Zmirou-Navier D, Deguen S (2013). A statistical procedure to create a neighborhood socioeconomic index for health inequalities analysis. *International Journal for Equity in Health*, **12**(1), 21.